

## SEQUENCE LISTING

### (1) GENERAL INFORMATION:

- (i) APPLICANT: NATSUKA, SHUNJI  
GERSTEN, KEVIN M.  
LOWE, JOHN B.
- (ii) TITLE OF INVENTION: MURINE ALPHA (1,3) FUCOSYLTRANSFERASE  
FUC-TVII, DNA ENCODING THE SAME, METHOD FOR PREPARING THE  
SAME, ANTIBODIES RECOGNIZING THE SAME, IMMUNOASSAYS FOR  
DETECTING THE SAME, PLASMIDS CONTAINING SUCH DNA
- (iii) NUMBER OF SEQUENCES: 4
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
P.C.
  - (B) STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
  - (C) CITY: ARLINGTON
  - (D) STATE: VA
  - (E) COUNTRY: USA
  - (F) ZIP: 22202
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/613,098
  - (B) FILING DATE: 08-MAR-1996
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: LAVALLEYE, JEAN-PAUL
  - (B) REGISTRATION NUMBER: 31,451
  - (C) REFERENCE/DOCKET NUMBER: 2363-114-55
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 703-413-3000
  - (B) TELEFAX: 703-413-2220

### (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 3594 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ACAAACAGGA AGGACAGCAG GCTCTGGCAG CCAGAAGCCT GTGGCCCCAA GCTGGCAGGA	60
TGGCCCCCTT CCTGCAGGTC CCCACAGCC TTCTGGGTTC CTGACACGAG AGAAGAGGTG	120
GGGCGGGGTG AAGTGAATC TGAAGCCAAA ATGTGACTCT CCTGGGGTCA CCAGCTTGGG	180
GAGAGGTGAA GAAAGATGCC GGGGCGGAAA CAAAGGGGCA GATATCACTA TGGTTATCTT	240
ACTAAGCACA GAGTAACTGA AAAAGCAAGG GTACCGCTGC CCACCTCGTG CCCACCTTAC	300
GTTATACCTC AAACCAGCTA GATAGTTTCT GATGGCACCC ATACCCTCCC TTCCCCTTTA	360
GGCATTGCGC AAGCTCTCCA CCACAATCTG GAAGTTATAC CCTGCGAGGG GATGGGCAGG	420
GCACTTCTGA GGTGCCAATC AGCCTGCACT CGCCTCTGCC CTGGCCATGG CACTGCTGTC	480
AGTTTCTTGG TACCTGTCTC AACAGCAGCC TTGTCACGTG AGACTATGGC TGGCGGTGGG	540
GGTGGGGGCA GGAATCCTAG AAGCACAGGA GTGACATAGG GTCGGGTCGG GCAGAGCGAA	600
GTGTAGGAGG TGATCCCCAA AGGGATGCTG GGGACGATCT GGCCAACACT GTCCTCCCAT	660
TCAAAACTCC CAGTCTGGAG CTCTGGGACA TGGACAAGCC AGGCCTGCTA TTCTCCATAC	720
AGGGCTCCAT AGTGTCTGGC TCAGCAGAGT GGGGGATCTG GTGGGGATGG AGGAAGCTTA	780
GCTAAAAGCT TTGTATAGGC TGAAGCTCTG AGTGACCCTG CTGGGCCACC CTACCCTGGT	840
CTGGGCTGGG TCATTGCATC CCCAGATTGG AAGGCTTGGT GAGATGGAGA GGAACCTTGG	900
CTACAAGCTA TAGCTTTGCC CACCAGAGCC TGCTGGAGGG GAATCAAACA AGCCTGGACC	960
TGAGGCTGGG ACTAGCTTTC CTGTTTCTGG AGTGGATGCC AACCCCCTGC CCACCAGCCT	1020
GCCTGTCCAC GCCAGGGACA CACAGACTCC TTCCCTTTCC AGACTGGAAA GCCCCCTCCT	1080
GGGAGAGCAG GAAGGAAGCA ACCTGCAACT CTTCCAGCCC TGGACCTTGG GCTGAACCTA	1140
CAGTTCAAGG TTTGTATGCT CACAGGTCTT GGCAGGGAAA GATAAGAATC CCCAGGGCAC	1200
CCTCCCCCCC GCCCCCAGT CCACTGCAGG TAGCTCCTGG GTCTGCCCTT CAGGGCAAGT	1260
GCTGACGCTC CATCAGACTG TGATGGGGCC CTTTTCTGAG GATGACAATT CTGAGAACAA	1320
GGCATTTTTTC TAGAGGTGGC AGAACAGCAT TTTGTGATGC CCGAGGATCT GGGAGCACAG	1380
GTCCAGCTTA ATGAGGGATT GGAGGAAGTG GGTATCATCA TTACAGGGAG GGGCCTCTGT	1440
GGCCTCCTGG GAAAATGCAG TTGCTCTCTT TGGGTGGCCT GGGGTTGTGT GGTGGGCAGA	1500
GGACGGAGGT GCTCATTGGG GGAAGGGATC ACTTCTGCTC AGAGTGCTCG CAAGGGCCTT	1560
TCCTTTTCCT GAAGGCAAGC AGGCCTCCTC CTCCTCCTCT TCCTCCTTCT CCTCTTCCTC	1620
CTCTTTCTCC ATATGCCTAG CTGGTCATTT CTAGGGACCA GCATGGTTGG GAAGGGGGCC	1680

TTGTCTTGGC	CTTCCTCTTG	TCTCAATTCC	CTCTTTGAGC	AGAAGACGGG	GTGGGTGGGG	1740
TAGGATTGGA	TAGTGGTTGA	TGCCAAAGAT	TGAAGGGGTA	GGGCGGGGCA	GAAGTGGGAA	1800
GGTCCCTGGC	TTCCTCACCT	TGGTAGATGG	TGAGGAGCCC	CAGAGGTTGA	GCTGAGCAGC	1860
AGCTGTGATT	TCAGGGTGCC	TCTGTTGGAG	AGGCTGCTGT	GATTTGAAAA	TCTTCTTTCC	1920
TTGGTGACAA	TTCCAGAAGG	CTCCAGATGA	ATTGTATTGG	TGAGTGCCTG	GCCCTTAAGC	1980
AGTCCCAGCT	GGGGATGATG	GGGATTTATG	GGTGTCCCTG	AGCCTAGGGT	GACAGGGCCT	2040
CTCCTTTTTT	TTTTATTCTG	CTTCAGGGTA	CCACCCACCC	AGGAGGCTGC	GGGCCTGGGG	2100
CGGCCTAGCT	GGAGGAGCAA	CATTCATGGT	AATTTGGTTT	TTCTGGCTGT	GGGGATCAGC	2160
TCCTGGAAGT	GCCCCGTGTC	CTCAGTCCAC	ACTCACCATC	CTTATCTGGC	ACTGGCCTTT	2220
CACCAACCGG	CCGCCAGAGC	TACCTGGTGA	CACCTGCACT	CGCTATGGCA	TGGCCAGCTG	2280
CCGTCTGAGT	GCTAACCGGA	GCCTGCTAGC	CAGTGCTGAT	GCTGTGGTCT	TCCACCACCG	2340
TGAGCTGCAA	ACCCGGCAAT	CTCTCCTACC	CCTGGACCAG	AGGCCACACG	GACAGCCTTG	2400
GGTCTGGGCC	TCCATGGAAT	CGCCCAGTAA	TACCCATGGT	CTCCATCGCT	TCCGGGGCAT	2460
CTTCAACTGG	GTGCTGAGCT	ATCGGCGTGA	TTCAGATATC	TTTGTACCCT	ACGGTCGCTT	2520
GGAGCCTCTC	TCTGGGCCCA	CATCCCCACT	ACCGGCCAAA	AGCAGGATGG	CTGCCTGGGT	2580
GATCAGCAAT	TTCCAGGAGC	GGCAGCAGCG	TGCAAAGCTG	TACCGGCAGC	TGGCCCCCTCA	2640
TCTGCAGGTG	GATGTGTTCG	GTCGCGCCAG	CGGACGGCCC	CTATGCGCTA	ATTGTCTGCT	2700
GCCCACTTTG	GCCCCGTACC	GCTTCTACCT	GGCCTTTGAG	AACTCACAGC	ATCGGGACTA	2760
CATCACTGAG	AAGTTCTGGC	GCAATGCCCT	GGCGGCTGGT	GCTGTACCCG	TGGCGCTGGG	2820
ACCTCCTCGG	GCCACCTACG	AGGCTTTTGT	GCCACCAGAT	GCCTTTGTAC	ACGTGGACGA	2880
CTTCAGCTCT	GCCCCGTGAAC	TGGCTGTCTT	CCTCGTCAGC	ATGAATGAGA	GTCGTTATCG	2940
TGGCTTCTTT	GCTTGGCGAG	ACCGGCTCCG	TGTGCGGCTC	CTGGGTGACT	GGAGGGAGCG	3000
CTTCTGCACC	ATCTGTGCCC	GCTACCCTTA	CTTGCCCCGC	AGCCAGGTCT	ATGAAGACCT	3060
TGAAAGCTGG	TTCCAGGCTT	GAATCCTGC	TGCTGGGAGA	GGCTGGATGG	GTGGGAGACT	3120
GATGTTGAAA	CCAAAGAGCT	GGGCATCCAG	GCTTTTGGTC	ACCATGGCAC	TACCCCAAGG	3180
CTTTTCCTGT	TCAGTGAGCA	GGAATTCAGG	ATATAAGGAG	AAGACTGGGC	TGAGATACCC	3240
TGGTGGGCTT	TAGAGTAGGG	GCCCAGGATA	AGAGACAATG	AATTAATGAG	GAGCATATGG	3300
GGAAGGTGGC	TGAGGGTCCC	TGACTTACCT	TGACCCATGG	CTGAAGGCTC	CATGCCCATG	3360
GCTGGAGCTG	GGACCCTACA	CTTCTATAGT	CAAGGTGCTT	AGCCTCAAGG	TTGCAGATGC	3420

ACCCTCTAGT ACTCTGGGTG CAGACTGTAC ACTGGGCGCA GGGGGTTGTG GAAGGACAGT	3480
GCAGATGATT CTGGGCTTTT GACACCACAG TTCCCCCAGG GAAAGAGGCA CTACTAATAA	3540
AAACACTGAC AGAAATCTCC TGGTCAAGTC TGTTAGGCAG CAGAGCTCGA ATTC	3594

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 393 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Pro	Thr	Pro	Cys	Pro	Pro	Ala	Cys	Leu	Ser	Thr	Pro	Gly	Thr	His	1	5	10	15
Arg	Leu	Leu	Pro	Phe	Pro	Asp	Trp	Lys	Ala	Pro	Ser	Trp	Glu	Ser	Arg	20	25	30	
Lys	Glu	Ala	Thr	Cys	Asn	Ser	Ser	Ser	Pro	Gly	Pro	Trp	Ala	Glu	Pro	35	40	45	
Thr	Val	Gln	Met	Asn	Cys	Ile	Gly	Tyr	His	Pro	Thr	Arg	Arg	Leu	Arg	50	55	60	
Ala	Trp	Gly	Gly	Leu	Ala	Gly	Gly	Ala	Thr	Phe	Met	Val	Ile	Trp	Phe	65	70	75	80
Phe	Trp	Leu	Trp	Gly	Ser	Ala	Pro	Gly	Ser	Ala	Pro	Val	Pro	Gln	Ser	85	90	95	
Thr	Leu	Thr	Ile	Leu	Ile	Trp	His	Trp	Pro	Phe	Thr	Asn	Arg	Pro	Pro	100	105	110	
Glu	Leu	Pro	Gly	Asp	Thr	Cys	Thr	Arg	Tyr	Gly	Met	Ala	Ser	Cys	Arg	115	120	125	
Leu	Ser	Ala	Asn	Arg	Ser	Leu	Leu	Ala	Ser	Ala	Asp	Ala	Val	Val	Phe	130	135	140	
His	His	Arg	Glu	Leu	Gln	Thr	Arg	Gln	Ser	Leu	Leu	Pro	Leu	Asp	Gln	145	150	155	160
Arg	Pro	His	Gly	Gln	Pro	Trp	Val	Trp	Ala	Ser	Met	Glu	Ser	Pro	Ser	165	170	175	
Asn	Thr	His	Gly	Leu	His	Arg	Phe	Arg	Gly	Ile	Phe	Asn	Trp	Val	Leu	180	185	190	

Ser	Tyr	Arg	Arg	Asp	Ser	Asp	Ile	Phe	Val	Pro	Tyr	Gly	Arg	Leu	Glu
		195					200					205			
Pro	Leu	Ser	Gly	Pro	Thr	Ser	Pro	Leu	Pro	Ala	Lys	Ser	Arg	Met	Ala
	210					215					220				
Ala	Trp	Val	Ile	Ser	Asn	Phe	Gln	Glu	Arg	Gln	Gln	Arg	Ala	Lys	Leu
225					230					235					240
Tyr	Arg	Gln	Leu	Ala	Pro	His	Leu	Gln	Val	Asp	Val	Phe	Gly	Arg	Ala
				245					250					255	
Ser	Gly	Arg	Pro	Leu	Cys	Ala	Asn	Cys	Leu	Leu	Pro	Thr	Leu	Ala	Arg
			260					265					270		
Tyr	Arg	Phe	Tyr	Leu	Ala	Phe	Glu	Asn	Ser	Gln	His	Arg	Asp	Tyr	Ile
		275					280					285			
Thr	Glu	Lys	Phe	Trp	Arg	Asn	Ala	Leu	Ala	Ala	Gly	Ala	Val	Pro	Val
	290					295					300				
Ala	Leu	Gly	Pro	Pro	Arg	Ala	Thr	Tyr	Glu	Ala	Phe	Val	Pro	Pro	Asp
305					310					315					320
Ala	Phe	Val	His	Val	Asp	Asp	Phe	Ser	Ser	Ala	Arg	Glu	Leu	Ala	Val
				325					330					335	
Phe	Leu	Val	Ser	Met	Asn	Glu	Ser	Arg	Tyr	Arg	Gly	Phe	Phe	Ala	Trp
			340					345					350		
Arg	Asp	Arg	Leu	Arg	Val	Arg	Leu	Leu	Gly	Asp	Trp	Arg	Glu	Arg	Phe
		355					360					365			
Cys	Thr	Ile	Cys	Ala	Arg	Tyr	Pro	Tyr	Leu	Pro	Arg	Ser	Gln	Val	Tyr
	370					375					380				
Glu	Asp	Leu	Glu	Ser	Trp	Phe	Gln	Ala							
385					390										

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCGCGGATCC CACCATCCTT ATCTGGCACT GGCCTTTCAC C

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 44 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid  
 (A) DESCRIPTION: /desc = "SYNTHETIC PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GCGCGGATCC AGTTCAAGCC TGGAACCAGC TTTCAAGGTC CTTC